

ESD ACCESSION LIST

DRI Call No. 78685Copy No. 1 of 2 cys.

ESD-TR-72-284, [REDACTED]

MTR-2403, Vol. I

SDBP — A SMALL DATA BASE PACKAGE

W. Amory

DECEMBER 1972

Prepared for

DEPUTY FOR COMMAND AND MANAGEMENT SYSTEMS

ELECTRONIC SYSTEMS DIVISION

AIR FORCE SYSTEMS COMMAND

UNITED STATES AIR FORCE

L. G. Hanscom Field, Bedford, Massachusetts



ESD RECORD COPY
RETURN TO
SCIENTIFIC & TECHNICAL INFORMATION DIVISION
(DRI), Building 1435

Approved for public release;
distribution unlimited.

Project 515A

Prepared by

THE MITRE CORPORATION

Bedford, Massachusetts

Contract F19628-71-C-0002

AD760103

When U.S. Government drawings, specifications, or other data are used for any purpose other than a definitely related government procurement operation, the government thereby incurs no responsibility nor any obligation whatsoever; and the fact that the government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data is not to be regarded by implication or otherwise, as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

Do not return this copy. Retain or destroy.

SDBP — A SMALL DATA BASE PACKAGE

W. Amory

DECEMBER 1972

Prepared for

DEPUTY FOR COMMAND AND MANAGEMENT SYSTEMS

ELECTRONIC SYSTEMS DIVISION

AIR FORCE SYSTEMS COMMAND

UNITED STATES AIR FORCE

L. G. Hanscom Field, Bedford, Massachusetts



Approved for public release;
distribution unlimited.

Project 515A

Prepared by

THE MITRE CORPORATION

Bedford, Massachusetts

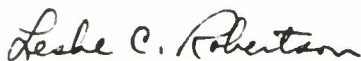
Contract F19628-71-C-0002

FOREWORD

The work described in this report was carried out under sponsorship of the Deputy for Command and Management Systems, Project 515A, by The MITRE Corporation, Bedford, Mass. , under Contract F19628-71-C-0002.

REVIEW AND APPROVAL

Publication of this technical report does not constitute Air Force approval of the report's findings or conclusions. It is published only for the exchange and stimulation of ideas.



LESLIE C. ROBERTSON, Colonel, USAF
Director of Systems Management
Deputy for Command and Management Systems

ABSTRACT

A collection of computer programs, documentation, procedures and data base design criteria have been packaged for handling small to medium-sized data bases. This first volume of the series gives a system overview of the resulting data base tools and techniques, which were developed largely in support of the MACIMS analysis effort. Subsequent volumes provide user information for each part of the system.

ACKNOWLEDGEMENT

Many people have contributed to the concepts and tools of the Small Data Base Package. I would particularly like to thank Sharon Hogan, Marilyn Anderson, Willie Johnson, and Nina Hanks for their help in implementing, testing and documenting the system, Mary Warford for her help with the artwork, and Elaine Valcourt and Bobbie Statkus for their typing and proofreading of these documents.

TABLE OF CONTENTS

	<u>Page</u>
LIST OF ILLUSTRATIONS	vi
LIST OF TABLES	vi
SECTION I INTRODUCTION	1
SECTION II FUNCTIONAL DISCUSSION	2
OVERVIEW	2
RAW DATA PROCESSING	2
FILE GENERATION AND UPDATE	5
FILE REPORTING	5
DATA DOCUMENTATION	5
SECTION III TECHNICAL CHARACTERISTICS	8
SECTION IV SAMPLE APPLICATION	9
APPENDIX DEFINITIONS AND DISCUSSIONS OF TERMS	17
INTRODUCTION	17
DEFINITIONS	17

LIST OF ILLUSTRATIONS

<u>Figure Number</u>		<u>Page</u>
1	Elements and Interfaces of a Data Base Project	3
2	Raw Data Transformation Options	4
3	File Generation and Updating	6
4	File Reporting Options	7
5	Sample Terminal Network Schematic	13
6	Terminal Network Files	14

LIST OF TABLES

<u>Table Number</u>		<u>Page</u>
I	Terminal Type File Format	9
II	Terminal Unit File Format	10
III	Terminal Link File Format	10
IV	Terminal Fixed-Format Card Formats	11
V	Terminal Unit Fixed-Format Card Format	12
VI	Terminal Link Fixed-Format Card Format	15

SECTION I

INTRODUCTION

The software system and concepts described by this series of documents were largely developed in support of the MACIMS project, where data on the MACIMS system was collected, organized, and analyzed in support of the MACIMS requirements development. This first volume of the series presents a system overview. Other volumes projected for the series include:

Volume II - KFFU (Keyed Flat File Update)

Volume III - KFFM (Keyed Flat File Merge)

Volume IV - FFRG (Flat File Report Generator)

Volume V - FFSP (Flat File Special Processor)

Volume VI - UNSTR (Unformatted String Text Receiver)

Volume VII - Data Base Documentation Tools

DS/2, another system associated with SDBP, is documented as DS/2, Users Manual, OS Version, November 1971, available from the System Development Corporation, Santa Monica, California.

Provision is made within the Small Data Base Package for data collection, generation and maintenance of data files, querying and reporting of data, and documentation of data elements and file structures.

SECTION II

FUNCTIONAL DISCUSSION

OVERVIEW

Figure 1, Elements and Interfaces of a Data Base Project, illustrates the major parts of a data base project and how they interface with each other. Each rectangular area of Figure 1 represents a large element, a common boundary between two blocks represents an interfacing activity, and the relative size of the boundaries represents the relative degrees of activity.

Data structures are kept simple for the Small Data Base Package. Files, known as "flat files," have records which, for a given file, are the same length and same format; that is, data for an item will occupy the same position in each record.

Data representation is simple also; all data in files is represented in character form with individual item values limited to 60 characters. While this can cause some expansion of file storage requirements, it also permits the use of commercially available systems such as DS/2 or SPSS for interactive querying or other types of processing, and encourages transferability of data bases between different types of computers.

RAW DATA PROCESSING

Figure 2, Raw Data Transformation Options, shows the flow of data from outside the system to fixed card formats. Three options are available. One option involves manual preparation of fixed format cards directly from raw data. A second option is to keypunch raw data onto free format cards and process the cards with the computer program UNSTR to automatically generate fixed format cards. The third is to transfer the raw data to paper tape (for example, by ASR35 Teletype), copy the paper tape to ASCII magnetic tape with a minicomputer such as D73's Interdata 3 computer, and then process this magnetic tape with UNSTR to automatically generate fixed format cards. In all cases, the purpose is ultimately to transfer free format raw data to fixed format cards for further processing by other elements of the system. Data validation will normally be done by special PL/1 programs designed for formats and values peculiar to individual files.

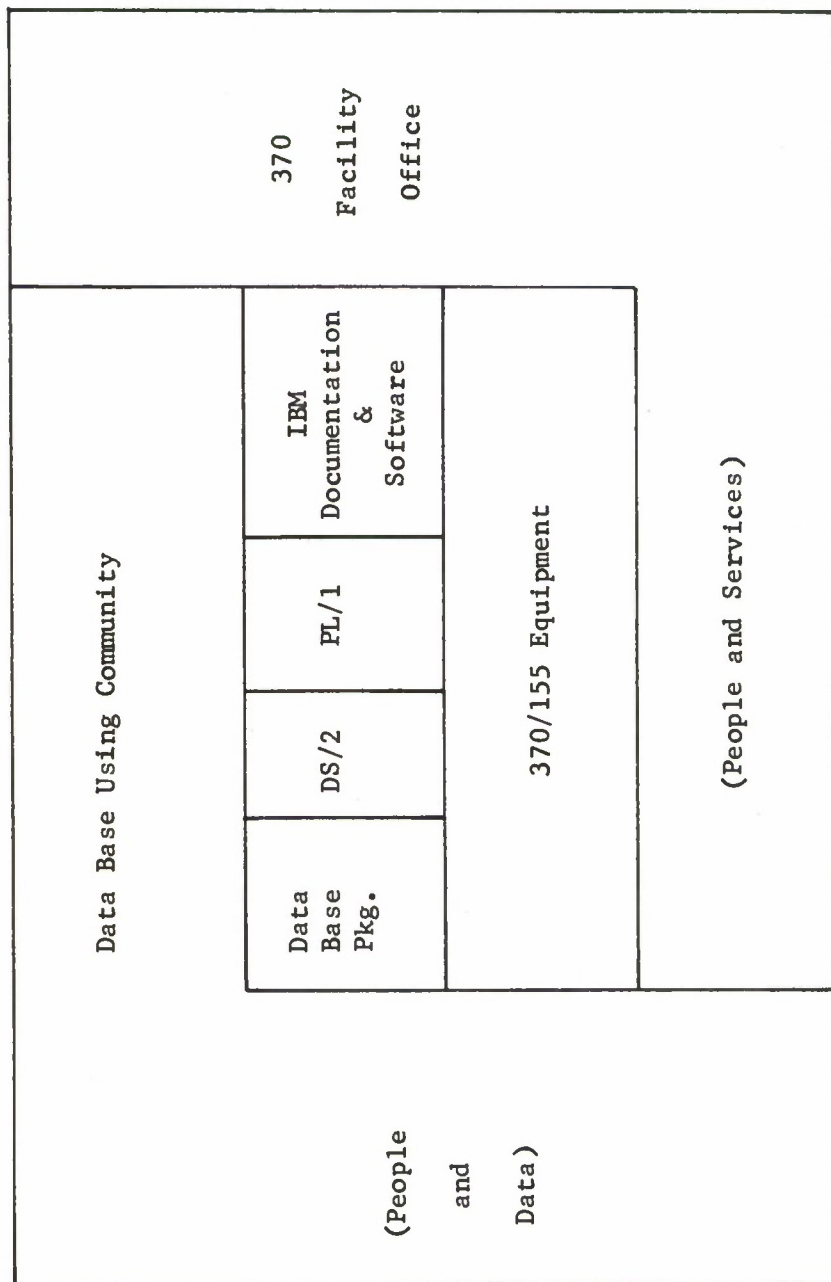


Figure 1. Elements and Interfaces of a Data Base Project

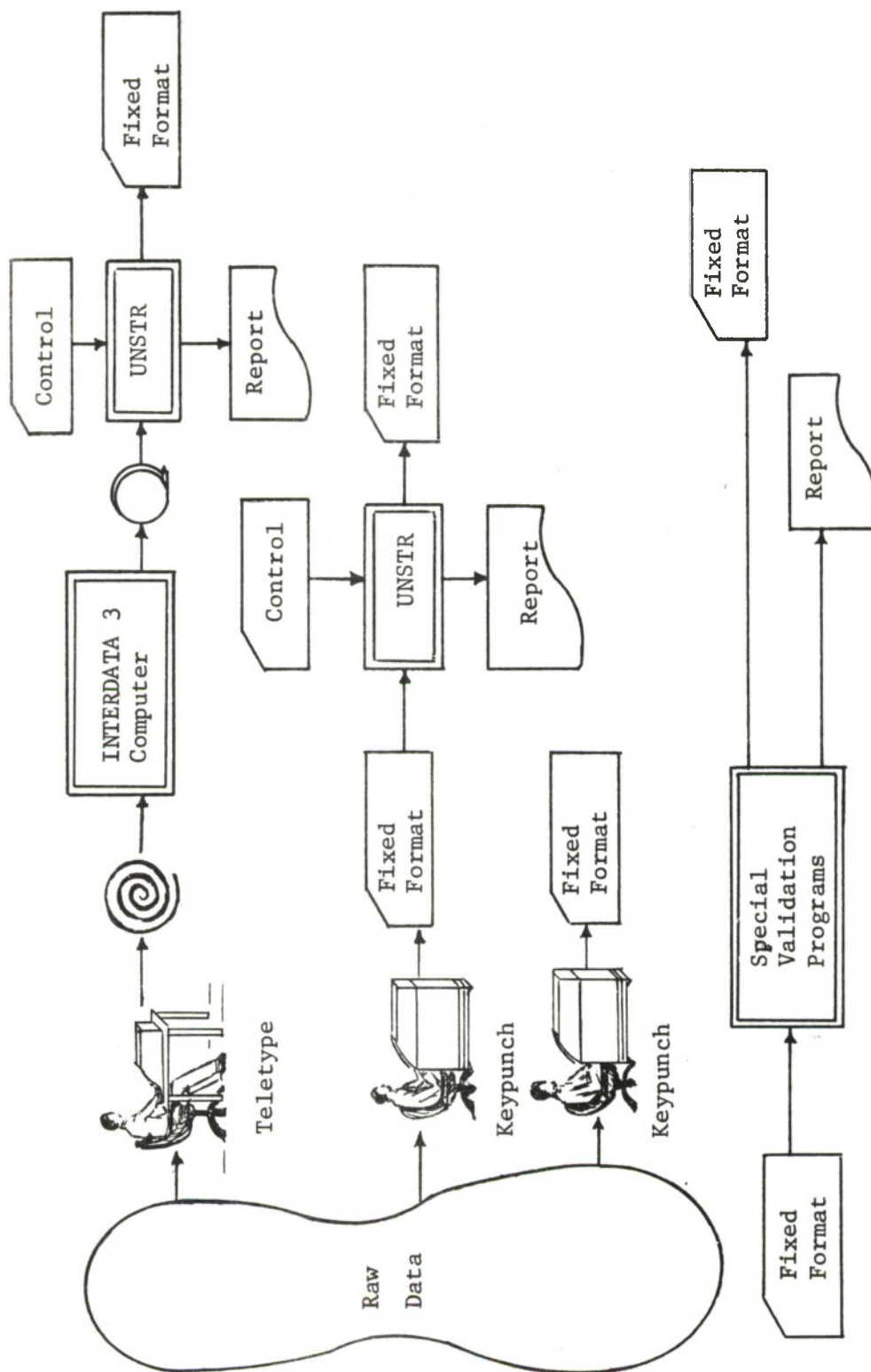


Figure 2. Raw Data Transformation Options

FILE GENERATION AND UPDATE

Figure 3, File Generation and Updating, shows how files are generated and updated from combinations of fixed format cards and files. Both operations may be performed with the update program, KFFU.

Updating involves fixed format cards, control information, and an input file. Updating includes deletion of individual data or of entire records, addition of individual data or of entire records, and modification of individual data.

For wholesale manipulation of data, as when an entire set of values for item is to be transformed into another set, the computer program FFSP provides a framework into which special subprograms can be inserted for manipulating the data.

FILE REPORTING

Figure 4, File Reporting Options, shows the file reporting process. Two processes are actually involved -- query reports and standard reports. An online query report capability is available with DS/2, an SDC query package. Standard reports may be produced with FFRG.

DATA DOCUMENTATION

Even modest-sized data bases have their problems with ensuring common understanding among users of the meaning of items and their format on cards and files. The data base documentation package provides for centralization of this kind of information with computer-generated reports, based on key-punched descriptions of meanings, locations, sizes and relationships of data items.

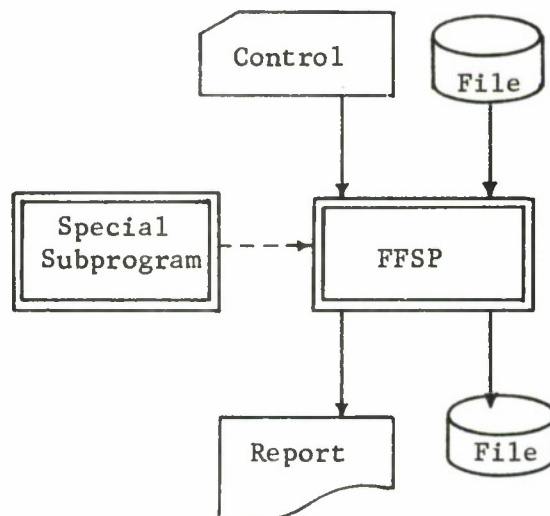
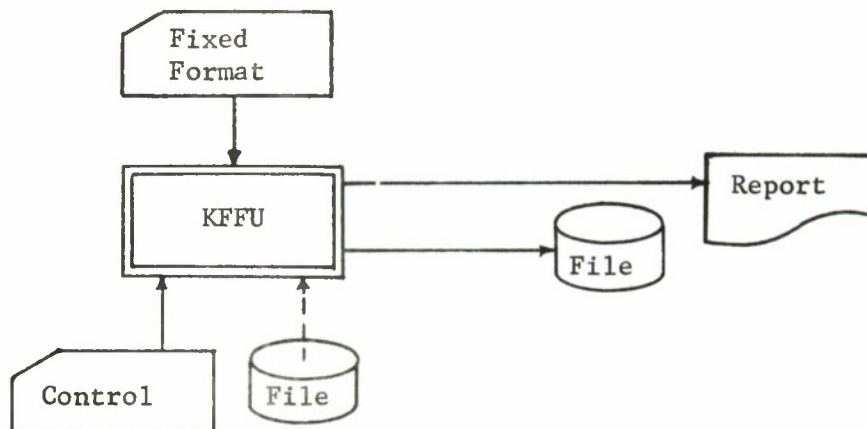


Figure 3. File Generation and Updating

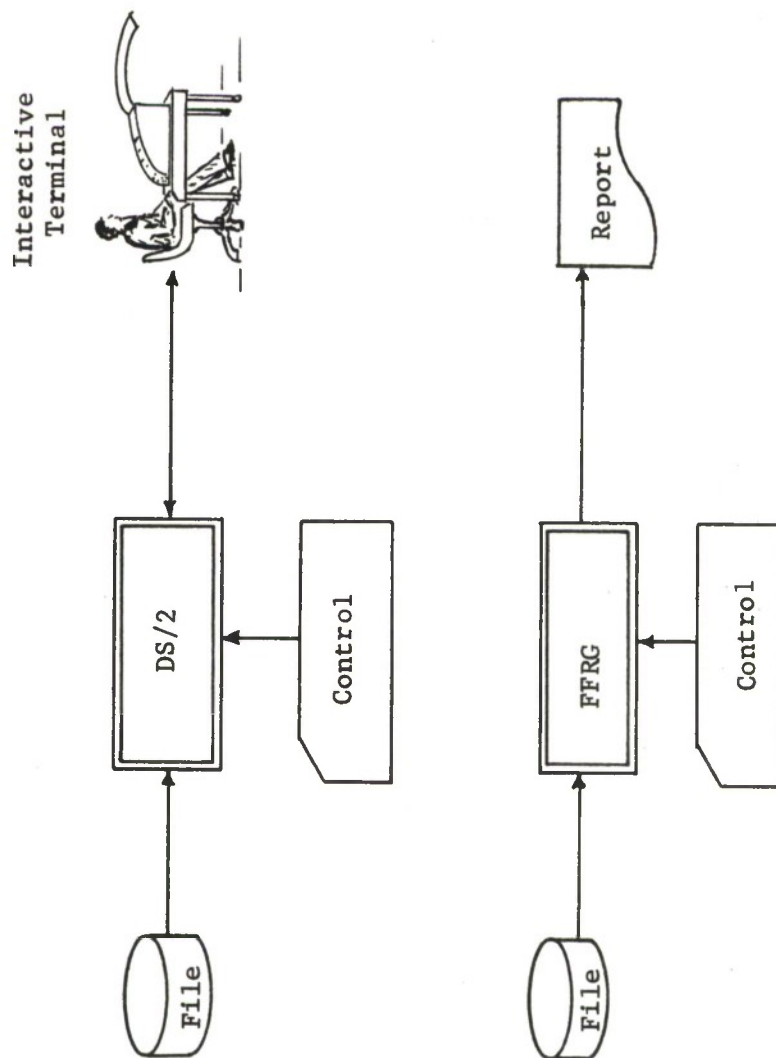


Figure 4. File Reporting Options

SECTION III

TECHNICAL CHARACTERISTICS

The data base package was written in PL/1 for an IBM 360 or 370. Partition storage requirements, which will vary somewhat between applications, will not normally exceed 128K.

Control information processed by members of the package is generally in free format, data-directed form (e.g., FRL=60 NCC(5) = 3). A few programs used fixed format control (card) inputs.

When no value is supplied or when a supplied value is illegal, an item will have blanks for its current value.

Item values are always character data, and may not exceed 60 characters in length.

SECTION IV

SAMPLE APPLICATION

To illustrate the use of the Small Data Base Package, a sample terminal network system is described below.

The system consists of three elements: terminals, locations, and links.

The data base consists of three files: the Terminal Type File, the Terminal Unit File, and the Terminal Link File.

The Terminal Type File contains data which distinguishes one type of terminal from another; the file is described in Table 1.

Table I

Terminal Type File Format

<u>Data Description</u>	<u>Code Name</u>	<u>Start Char</u>	<u>Number Chars</u>
Terminal Type Code - (alphanumerics identifying the type--e.g., 33ASR)	TC	2	5
Terminal Type Name - (alphanumerics, left adjusted)	TN	7	12
Terminal I/O Device Description - (alphanumerics, left adjusted)	DD	19	27
Terminal I/O Capability Description - (alphanumerics, left adjusted)	CD	46	24
Number of Terminal Types - (digits, leading zero)	NT	70	2
Transfer Rate for Terminal - (digits, baud)	TR	72	4

The Terminal Unit File contains data which defines each terminal; the file is described in Table 2.

Table II

Terminal Unit File Format

<u>Data Description</u>	<u>Code Name</u>	<u>Start Char</u>	<u>Number Chars</u>
Terminal Unit Code - (alpha- numerics identifying the unit--e.g., 3288A)	UC	2	5
Terminal Type Code - (alpha- numerics identifying the type--e.g., 33ASR)	TC	7	5
Corporate Code - (alphanumeric identifying the site, left adjusted)	CC	12	15
Building Number - (digits, leading zero)	BN	27	2
Room Code - (alphanumeric-- e.g., 1B245)	RC	29	4
Start Time Availability - (Time SA of Day when terminal first available--e.g., 0800)		33	4
End Time Availability - (Time EA of Day after which terminal not available--e.g., 2200)	EA	37	4

The Terminal Link File contains data which defines each line of the terminal network; the file is described in Table 3.

Table III

Terminal Link File Format

<u>Data Description</u>	<u>Code Name</u>	<u>Start Char</u>	<u>Number Chars</u>
Link Code - (alphanumeric identifier of communications link-- e.g., LINK1)	LC	1	5
Terminal Unit Code - (alphanumeric identifying the unit-- e.g., 3288A)	UC	7	5
I/O Code - (I=input, O=output)	IC	12	1
Transfer Rate for Link - (digits, TR line rate baud)		13	4

Certain code items are used as keys to relate data in one file with that of another file. For example, to find the Terminal Transfer rate of a unit described in the Terminal Unit File, the Terminal

Type Code value of that entry is matched to the Terminal Type Code value of entries in the Terminal Type File. The transfer rate associated with the matched code is the correct rate for the unit (e.g., unit 5180K has the transfer rate of 10 baud).

Figure 5, Sample Terminal Network Schematic, describes the system as a whole in schematic fashion.

Figure 6, Terminal Network Files, describes sample versions of the three files.

Tables 4-6 describe fixed-format card formats for all three files; Figure 7 lists cards which would produce the files of Figure 6.

Table IV

Terminal Fixed-Format Card Formats

<u>Card Type A</u>			
<u>Item Code</u>	<u>Start Col</u>	<u>Number Cols</u>	<u>Comments</u>
-	2	2	Contains TU to identify file
-	4	1	Contains A to identify card type
TC	6	5	Terminal Type Code
TN	11	12	Terminal Type Name
DD	23	27	Terminal I/O Device Description

Table IV (Continued)

Card Type B

<u>Item Code</u>	<u>Start Col</u>	<u>Number Cols</u>	<u>Comments</u>
-	2	2	Contains TL to identify file
-	4	1	Contains B to identify card type
TC	6	5	Terminal Type Code
CD	11	24	Terminal I/O Capability Description
NT	35	2	Number of Terminal Types
TR	37	4	Transfer Rate for Terminal

Table V

Terminal Unit Fixed-Format Card Format

<u>Item Code</u>	<u>Start Col</u>	<u>Number Cols</u>	<u>Comments</u>
-	2	2	Contains TU to identify file
-	4	1	Contains A to identify card type
UC	6	5	Terminal Unit Code
CC	11	15	Corporate Code
BN	26	2	Building Number
RC	28	4	Room Code
SA	32	4	Start Time Availability
EA	36	4	End Time Availability

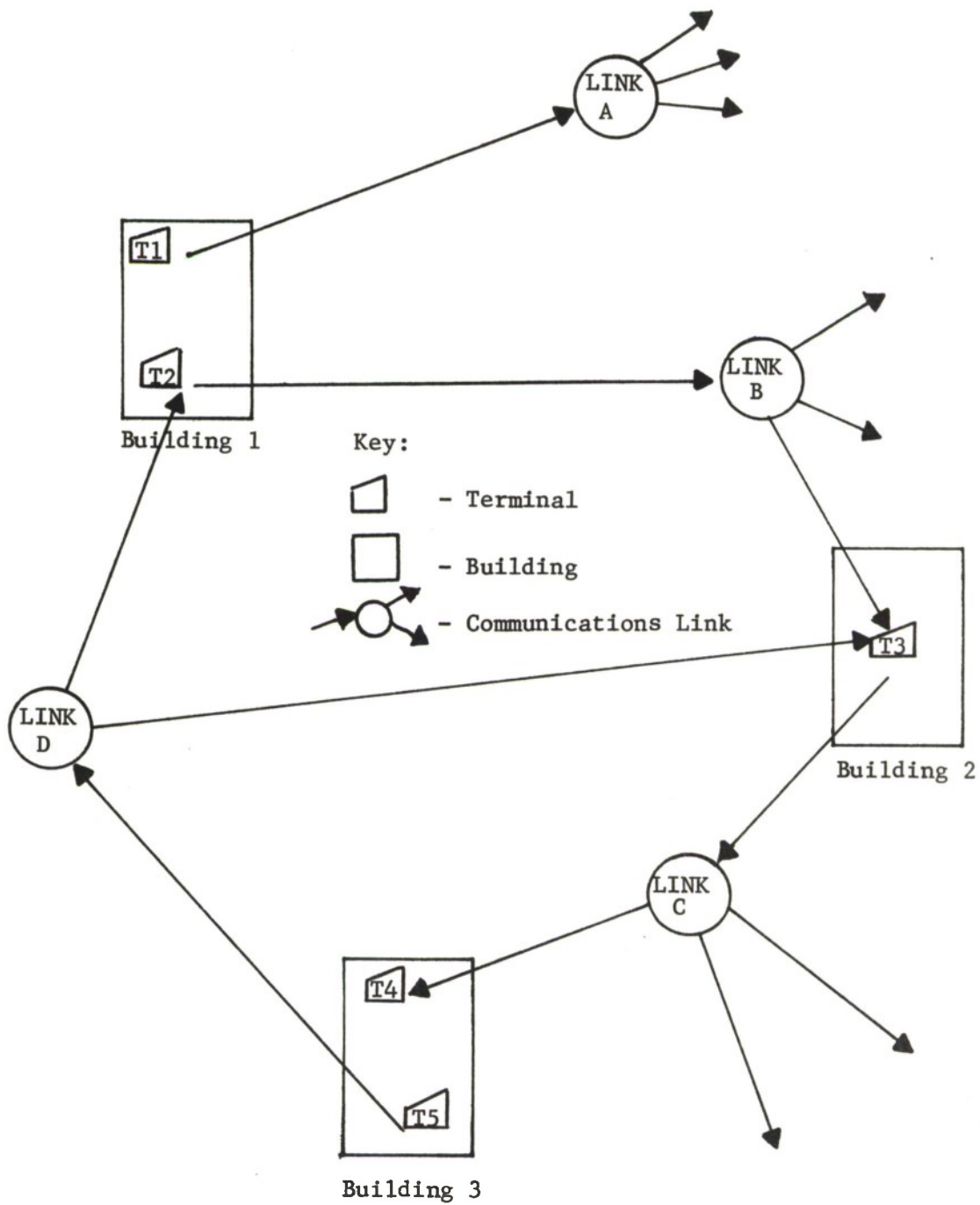


Figure 5. Sample Terminal Network Schematic

Terminal Type File

Data: 33ASRTELETYPE33PAPER3TAPE3PUNCH33READER3AUTOMATIC3SEND-RECEIVER30010
 35KSRTELEPRINTER3PAGE3PRINTER33RECEIVER33KEYBOARD3SEND-RECEIVER30015
 Item: TC TN DD CD NTTR

Terminal Unit File

Data: 3278A33ASRMITRE/BEDFORD3023E2407002400
 3288A33ASRMITRE/BEDFORD3031B0107002400
 3594A33ASRMITRE/BEDFORD3014B1707002400
 5179K35KSRMITRE/BEDFORD3014B1707002400
 5180K35KSRMITRE/BEDFORD3031B9807002400
 Item: UC TC CC BNRC SA LA

Terminal Link File

Data: 3594AI1800 Data: 35179K02400
 3278AI1800 3278A01800
 3288A01800 3288A01800
 3288AI1800 3288AI1800
 Item: LC UC TR LC UC TR IC

Note: blanks are represented by 3

Figure 6. Terminal Network Files

Table VI

Terminal Link Fixed-Format Card Format

<u>Item Code</u>	<u>Start Col</u>	<u>Number Cols</u>	<u>Comments</u>
-	2	2	Contains TL to identify file
-	4	1	Contains A to identify card type
LC	6	5	Link Code
UC	11	5	Terminal Unit Code
IC	16	1	I/O Code (I or O)
TR	17	4	Transfer Rate of Link (Band)

Using KFFU, the files described above can be generated from the fixed-format cards also described above. Reports can be generated with FFRG directly from these files or from merged temporary files generated with KFFM. All the control information needed by these programs is given in the above tables; the only data which appears in more than one file is key (code) data to permit cross-referencing.

APPENDIX

DEFINITIONS AND DISCUSSIONS OF TERMS

INTRODUCTION

This material serves two purposes. The first is to define what is meant by certain frequently-used terms in the context of the Small Data Base Package. The second is to present additional material on the concepts upon which the design and use of the Small Data Base Package is based. The definitions are ordered alphabetically to permit easy access for the first purpose; to use this material for understanding design and use concepts, however, the following reading order is recommended:

- Software
- Hardware
- Data
- Data Base
- File
- Flat File
- Keyed Flat File
- Element
- Entry
- Record
- Attribute
- Item
- Fixed Format Cards

DEFINITIONS

Attributes Abstract features of elements for which values can be determined and assigned to items as data. Items are the concrete (i.e., data base) equivalent of attributes. If, for example, terminals were an element of a transportation data base such as MACIMS', attributes could include character rate, model number, and type of display (CRT, printed paper, punched paper, etc.). These are all features of terminals which would have only one value for any one terminal. Some attributes may have more than one value, such as terminal location and terminal number. When the range of values is large enough (a decision by the data base designer), a separate file will be established. In this case, one file could be a Terminal

Characteristics file and the other a straight Terminal file; the two files would contain the model number of the terminal as an attribute, which could then be used as a key to associate data between the two files.

Data	The values assigned items in a data base. All blanks in a value indicates that the item has yet to be assigned a meaningful value.
Data Base	Data, in machine-accessible form (punched cards, magnetic tape, disc, etc.), organized as one or more files. The data may come from external sources (people, books, other data bases, etc.) or may be generated from other data in the same data base.
Element	An abstraction of the data base at the lowest level which is defined by one or more attributes. For example, people could be an element of a data base, with attributes such as height, weight, employee number, work category, and telephone number.
Entry	The collection of item values in a file which, taken together, define one of the file's elements (e.g., George Washington as distinct from Abraham Lincoln in a Presidential file).
FFRG	Flat File Report Generator, the part of the SDBP which provides for writing reports from flat files.
FFSP	Flat File Special Processor, the part of the SDBP which provides for special processing of flat files which would be difficult for KFFU (e.g., mass conversion of an item's values).
File	A subset of a data base which has been made on the basis of relationships between items of the data base. In the approach used with the Small Data Base Package, a file is organized around an element of the data base (such as a Personnel Characteristics File around people data). The data in a file represents values for attributes of the element (like a person's height) which have one value (or at most a very few values, which would then be assigned separate names like PHONE1, PHONE2) for any one element. For example, children's names in a Personnel data base would

probably be broken off into a separate file along with their individual attributes (such as birth date) and a key, which is an attribute to associate data from one file to another. In the case of a Children's File, the key might be the parent's employee number, which would also appear in the Personal Characteristics file entry for the parent.

**Fixed Format
Cards**

The standard data source (but not necessarily the only one) from which files are directly generated or updated. Fixed Format Cards have column 1 blank, columns 2-3 reserved for the file code (e.g., PC for Personal Characteristics), column 4 reserved for the card type, and column 5 blank. Column 6 is reserved for the start of the key data, which is repeated for every card type containing data for the same element. After the key data come attribute data, usually arranged so as to fit the greatest number of items on the fewest card types (an item cannot be continued across card types). The use of more than one card type becomes necessary when the net length of the items associated with elements of a file exceeds the room available for data on a card (75 columns including the key item).

Flat File

A file in which every entry has the same items (same size, same meaning, same relative position in the record, but not usually the same values) as every other entry. In the Small Data Base Package design approach, the first character of every record is reserved as a flag to indicate data which was new or changed when the file was last updated or generated.

Keyed Flat File

A flat file which has at least one piece of data (like an employee number) which can be used to relate to data in another keyed flat file (like a Children's File). The entries of a keyed flat file are sorted by key values. The key item will normally start in the second character of a file record and be followed by data.

KFFM

Keyed Flat File Merge, the part of the SDBP which provides for merging portions of two keyed flat files based on common key values.

KFFU	Keyed Flat File Update, the part of the SDBP which provides for updating (including generating and maintaining) keyed flat files.
Hardware	Equipment (such as computers, tapes, and disks) which support the operation of computer programs.
Item	The concrete equivalent of attributes. Items are assigned names (e.g., EMPNO) and standard sizes (e.g., 35 characters). For each file in which the item appears, the item will also have a starting position in the file (e.g., character 15); usually the size of the item in the file will be the same as the standard size. The item will normally also have a card type assignment (see Fixed Format Cards) and a card starting column for each file; the exception to this is the item which is generated from data in the same or other files.
Record	The hardware representation of Entry. A record is the collection of data associated with an element involved in each read or write operation directed to a file.
SDBP	The Small Data Base Package, a system of concepts, procedures and programs oriented towards the management of small- to medium-sized data bases.
Software	Computer programs, as well as documentation of computer programs and of data.

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) The MITRE Corporation Bedford, Mass. 01730		2a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED	
		2b. GROUP	
3. REPORT TITLE SDBP — A SMALL DATA BASE PACKAGE			
4. DESCRIPTIVE NOTES (Type of report and inclusive dates)			
5. AUTHOR(S) (First name, middle initial, last name) W. Amory			
6. REPORT DATE DECEMBER 1972		7a. TOTAL NO. OF PAGES 25	7b. NO. OF REFS 1
8a. CONTRACT OR GRANT NO. F19628-71-C-0002		9a. ORIGINATOR'S REPORT NUMBER(S) ESD-TR-72-284, [REDACTED]	
b. PROJECT NO. 515A			
c.		9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)	
d.		MTR-2403 Vol. I	
10. DISTRIBUTION STATEMENT Approved for public release; distribution unlimited.			
11. SUPPLEMENTARY NOTES		12. SPONSORING MILITARY ACTIVITY Deputy for Command and Management Systems, Electronic Systems Division (AFSC), L. G. Hanscom Field, Bedford, Mass. 01730	
13. ABSTRACT A collection of computer programs, documentation, procedures and data base design criteria have been packaged for handling small to medium-sized data bases. This first volume of the series gives a system overview of the resulting data base tools and techniques, which were developed largely in support of the MACIMS analysis effort. Subsequent volumes provide user information for each part of the system.			

